

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 04:51:45 / Search time 55 Seconds

(without alignments)
102.745 Million cell updates/sec

Title: US-10-054-488-1

Perfect score: 109
Sequence: 1 GVTSAPDTPRAPGSTAPPAH 20

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: A_Geneseq_29Jan04:*

1: Geneseqp1980s: *
2: Geneseqp1990s: *
3: Geneseqp2000s: *
4: Geneseqp2001s: *
5: Geneseqp2002s: *
6: Geneseqp2003as: *
7: Geneseqp2003bs: *
8: Geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	109	100.0	20	2	AAW27393 Human epi
2	109	100.0	20	2	AAW67589 T-cell ac
3	109	100.0	20	3	AAW80111 MUC-1 rep
4	109	100.0	20	3	AAW80915 MUC-1 der
5	109	100.0	20	5	AAW82063 Antigenic
6	109	100.0	20	6	AAW33943 MUC 1 pep
7	109	100.0	20	6	AAW30200 Human MUC
8	109	100.0	20	7	ADB84183 Human MUC
9	109	100.0	20	7	ADD88878 Human MUC
10	109	100.0	25	2	AAW72724 Mucin pep
11	109	100.0	28	4	AAW03361 Mucin tan
12	109	100.0	30	4	AAW72472 Fusion pr
13	109	100.0	31	4	AAW46087 Human MUC
14	109	100.0	36	5	AAW82082 T-cell sp
15	109	100.0	38	5	AAW82074 T-cell sp
16	109	100.0	40	2	AAW68002 Mucin rep
17	109	100.0	40	2	AAW54873 Carcinoma
18	109	100.0	40	5	AAW72703 Human muc
19	109	100.0	40	5	AAW56039 Repeat mo
20	109	100.0	41	6	AAW96170 MUC1 repe
21	109	100.0	43	6	AAW33852 Dipeptidic
22	109	100.0	43	7	ADD88870 Synthetic
23	109	100.0	46	6	AAW56035 Histidine
24	109	100.0	46	6	AAW33936 MUC 1 pep
25	109	100.0	50	2	AAW35739 Mucin pep

26	109	100.0	51	2	AAW31697 Mucin pep
27	109	100.0	100	5	AAW76181 Synthetic
28	109	100.0	105	2	AAW68022 Mucin pep
29	109	100.0	105	2	AAW72697 Synthetic
30	109	100.0	134	2	AAW29894 Human IP-
31	109	100.0	137	2	AAW29893 Human MCP
32	109	100.0	138	2	AAW29806 Murine IP
33	109	100.0	156	2	AAW29807 Murine MC
34	109	100.0	172	2	AAW29895 Human MDC
35	109	100.0	173	3	AAW71021 Human MUC
36	109	100.0	177	2	AAW29896 Human SDF
37	109	100.0	216	3	AAW92665 MUC-1 ana
38	109	100.0	295	3	AAW71027 Ubiquitin
39	109	100.0	455	3	AAW71024 Human Muc
40	109	100.0	475	4	AAW09508 Human muc
41	109	100.0	475	4	AAW00573 Human MUC
42	109	100.0	475	5	AAW7476 Human MUC
43	109	100.0	475	6	AAW50567 Mucin 1 (
44	109	100.0	475	6	AAW37800 Human muc
45	109	100.0	475	7	ADD14120 Human src

ALIGNMENTS

AAW27393	standard; peptide; 20 AA.
AAW27393:	
14-APR-1998 (first entry)	
Human epithelial cell mucin MUC1 modified epitope.	
Hepatitis B virus; HBV; HBCAg; tumour antigen epitope; chimeric;	
Ha-ras oncogene; epithelial cell mucin; MUC1; human; tumour cell growth.	
Homo sapiens.	
Key	Location/Qualifiers
Region	6..10 /note="immunogenic sequence"
WO9735008-A1.	
25-SEP-1997.	
21-MAR-1997;	97WO-US004656.
21-MAR-1996;	96US-0013839P.
(USSH) US DEPT HEALTH & HUMAN SERVICES.	
Kwak LW, Biragyn A;	
WPI; 1997-480214/44.	
Chimeric polypeptide comprising hepatitis B virus core antigen and tumour antigen epitope - useful to inhibit tumour cell growth.	
Example 1; Page 22; 50pp; English.	
This is a modified epitope of the human epithelial cell mucin MUC1. The immunogenic sequence is located internally in this modified MUC1 epitope. This is used in the construction of a chimeric polypeptide comprising a Hepatitis B virus core antigen (HBcAg) and a tumour antigen epitope. The tumour antigen epitope is from the human epithelial cell mucin or is from the Ha-ras oncogene. The polypeptide can be used to inhibit the growth of tumour cells, especially by inducing an immune response to generate antibodies against a pathological or harmful condition	
Sequence 20 AA;	

Query Match 100.0%; Score 109; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 7.7e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVTSA PDTRPAPGSTAPPAH 20
 DB 1 GVTSA PDTRPAPGSTAPPAH 20

RESULT 2

AAW67589
 ID AAW67589 standard; peptide; 20 AA.

AC AAW67589;

XX 02-MAR-1999 (first entry)

DE T-cell activation peptide #7.

XX Activated T helper cell; CD4+; cytotoxic T cell; CD8+; liposome; epitope;
 KM peripheral blood lymphocyte; antigen-presenting cell; APC; virus; tumour;
 KM bacterium; parasite; cytokine; vaccine; cancer; malaria; HIV; hepatitis;
 KM tuberculosis.

XX Synthetic.

XX WO9850527-A1.

XX 12-NOV-1998.

XX 07-MAY-1998; 98WO-US009288.

XX 08-MAY-1997; 97US-0045349P.

XX (BIOM-) BIOMIRA INC.

XX Agrawal B, Krantz MJ, Reddish MA, Longenecker BM;

XX WPI; 1999-034715/03.

PT Method of activation of T cells - by exposure to antigen-presenting cells
 PT loaded with antigen in liposome, used for, e.g. treating cancer and
 PT microbial infections.

XX Claim 14; Page 49; 75pp; English.

XX Peptides AAW67583-W67611 are used to produce activated T helper (CD4+) and cytotoxic (CD8+) T-cells. The activated T cells are produced by treating peripheral blood lymphocytes with liposome-encapsulated peptide antigen to generate Ag-loaded antigen-presenting cells (APC), contacting native or anergic T-cells with these APC, and isolating the resulting activated T-cells. The cells are specific for a particular antigen, particularly one derived from a tumour, but also those from viruses, bacteria and other parasites. It can also be used to identify antigens and epitopes able to generate an Ag-specific T-cell response (by assessing proliferation and cytokine release). Also the Ag-loaded APC can be used as cellular vaccines for treating cancer (claimed) or other diseases (e.g. malaria, human immune deficiency virus infection, hepatitis, tuberculosis). The activated T-cells can be used to treat the same conditions by adoptive T-cell transfer therapy

XX Sequence 20 AA;

Query Match 100.0%; Score 109; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 7.7e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVTSA PDTRPAPGSTAPPAH 20
 DB 1 GVTSA PDTRPAPGSTAPPAH 20

RESULT 3

AAW80111
 ID AAW80111 standard; peptide; 20 AA.

AC AAW80111;

XX 17-MAY-2000 (first entry)

DE MUC-1 repeat peptide sequence.

XX Human; MUC-1; detection; T-cell activation; mucin; anti-inflammatory;
 KM immunomodulator; antitumour; antidiabetic; antiallergic;
 KM dermatological; antidiabetic; nephrotoxic; antihypertensive; antianemic;
 KM neuroprotective; hepatoprotective; uropathic; ophthalmological; antiviral;
 KM cytostatic; autoimmune disorder; inflammatory disorder; viral disease;
 KM cancer.

XX Homo sapiens.

XX WO200000828-A1.

XX 06-JAN-2000.

XX 25-JUN-1999; 99WO-US012820.

XX 26-JUN-1998; 98US-0090916P.

XX (BIOM-) BIOMIRA INC.

XX Agrawal B, Longenecker BM;

XX WPI; 2000-170935/15.

PT Detecting T-cell activation by measuring the amount of MUC-1 expression
 PT useful for diagnosing or treating autoimmune or inflammatory disorders,
 PT viral disease or cancer.

XX Disclosure; Page 11; 40pp; English.

XX A method has been developed for detecting T-cell activation by evaluating
 CC the amount of MUC-1 mucin expression in a T-cell compared to a non-
 CC activated control. The method is useful for treating disorders associated
 CC with T-cell activation, using an agent (antibody/antagonist) that
 CC modulates MUC-1 activity. The T-cell activation associated disorders may
 CC be autoimmune or inflammatory disorders (e.g. inflammatory arthritis,
 CC rheumatoid arthritis, psoriasis, allergies, allergic contact dermatitis,
 CC ankylosing spondylitis, myasthenia gravis, systemic lupus erythematosus,
 CC polyarteritis nodosa, Goodpasture's syndrome, isopathic thrombocytopenic
 CC purpura, autoimmune haemolytic anaemia, Graves' disease, rheumatic fever,
 CC pernicious anaemia, insulin-resistant diabetes mellitus, bullous
 CC pemphigus vulgaris, viral myocarditis (Cockayne B virus response),
 CC autoimmune thyroiditis (Hashimoto's disease), male infertility
 CC (autoimmune), sarcoidosis, allergic encephalomyelitis, multiple
 CC sclerosis, Sjogren's disease, Reiter's disease, Celiac disease,
 CC symptomatic ophthalmia, and primary biliary cirrhosis), viral disease or
 CC cancer. The present sequence represents a MUC-1 peptide from the present
 CC invention

XX Sequence 20 AA;

Query Match 100.0%; Score 109; DB 3; Length 20;
 Best Local Similarity 100.0%; Pred. No. 7.7e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVTSA PDTRPAPGSTAPPAH 20
 DB 1 GVTSA PDTRPAPGSTAPPAH 20

RESULT 4

AAW09915
 ID AAW09915 standard; peptide; 20 AA.

XX

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 04:53:40 ; Search time 24 Seconds
(without alignments)
43.022 Million cell updates/sec

Title: US-10-054-488-1

Perfect score: 109

Sequence: 1 GVTSPDTRPAPGSTAPPAAH 20

Scoring table: BLOSUM62

Searched: Gapop 10.0, Gapext 0.5

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued Patents AA*
2: /cgn2_6/ptodata/2/1aa/5A COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/5B COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6A COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	109	100.0	20	3 US-09-339-944-6	Sequence 6, Appl
2	109	100.0	20	4 US-09-487-232-10	Sequence 10, Appl
3	109	100.0	20	4 US-09-651-265-6	Sequence 6, Appl
4	109	100.0	25	2 US-08-288-059-28	Sequence 28, Appl
5	109	100.0	30	3 US-08-134-198B-13	Sequence 13, Appl
6	109	100.0	40	1 US-08-099-354-1	Sequence 1, Appl
7	109	100.0	40	2 US-08-288-059-7	Sequence 7, Appl
8	109	100.0	134	4 US-09-646-028-1	Sequence 1, Appl
9	109	100.0	137	4 US-09-646-028-2	Sequence 2, Appl
10	109	100.0	138	4 US-09-646-028-3	Sequence 3, Appl
11	109	100.0	138	4 US-09-646-028-4	Sequence 4, Appl
12	109	100.0	172	4 US-09-646-028-19	Sequence 49, Appl
13	109	100.0	177	4 US-09-646-028-54	Sequence 54, Appl
14	109	94.5	20	1 US-08-328-536-1	Sequence 1, Appl
15	109	94.5	28	2 US-08-488-161-9	Sequence 9, Appl
16	109	94.5	28	3 US-09-273-685-9	Sequence 9, Appl
17	109	94.5	28	5 PCT-US95-11934-9	Sequence 9, Appl
18	109	92.7	25	4 US-09-217-1068-3	Sequence 9, Appl
19	99	90.8	20	4 US-09-487-232-11	Sequence 11, Appl
20	97	89.0	24	4 US-08-737-896-5	Sequence 5, Appl
21	97	89.0	24	4 US-09-497-232-23	Sequence 23, Appl
22	97	89.0	24	5 PCT-US96-09951-5	Sequence 5, Appl
23	97	89.0	25	4 US-09-487-232-5	Sequence 5, Appl
24	97	89.0	30	3 US-08-737-896-6	Sequence 6, Appl
25	97	89.0	30	5 PCT-US96-09951-6	Sequence 6, Appl
26	94	86.2	19	4 US-09-593-870A-22	Sequence 22, Appl
27	94	86.2	20	2 US-08-833-807-1	Sequence 1, Appl

28	94	86.2	20	3 US-09-223-043-1	Sequence 1, Appl
29	94	86.2	20	4 US-09-593-870A-1	Sequence 1, Appl
30	91	83.5	1867	2 US-08-479-537A-5	Sequence 5, Appl
31	91	83.5	1867	3 US-09-083-116-5	Sequence 5, Appl
32	91	83.5	1867	4 US-09-134-916A-5	Sequence 5, Appl
33	91	83.5	2035	2 US-08-479-537A-2	Sequence 2, Appl
34	91	83.5	2035	3 US-09-083-116-2	Sequence 2, Appl
35	91	83.5	2035	4 US-09-134-916A-2	Sequence 2, Appl
36	90	82.6	16	4 US-09-043-731-18	Sequence 18, Appl
37	90	82.6	20	2 US-08-833-807-8	Sequence 8, Appl
38	90	82.6	20	3 US-09-223-043-8	Sequence 8, Appl
39	90	82.6	20	3 US-09-221-351-1	Sequence 1, Appl
40	90	82.6	20	4 US-09-043-731-16	Sequence 16, Appl
41	90	82.6	20	4 US-09-593-870A-20	Sequence 20, Appl
42	90	82.6	21	2 US-08-833-807-7	Sequence 7, Appl
43	90	82.6	21	3 US-09-223-043-7	Sequence 7, Appl
44	90	82.6	21	4 US-09-043-731-15	Sequence 15, Appl
45	90	82.6	21	4 US-09-593-870A-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-09-339-944-6
Sequence 6, Application US/09339944

Patent No. 6114129
GENERAL INFORMATION:
APPLICANT: AGRAMAL, Babita
TITLE OF INVENTION: METHODS OF DETECTING T-CELL ACTIVATION AND TREATING
FILE REFERENCE: 042881/0129
CURRENT FILING DATE: 1999-06-25
EARLIER FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 20
TYPE: PRT
ORGANISM: Homo sapiens
US-09-339-944-6

Query Match 100.0%; Score 109; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVTSPDTRPAPGSTAPPAAH 20
Db 1 GVTSPDTRPAPGSTAPPAAH 20

RESULT 2
US-09-497-232-10
Sequence 10, Application US/09497232

Patent No. 6600012
GENERAL INFORMATION:
APPLICANT: AGRAMAL, Babita
KRANTZ, Mark J.
REDDISH, Mark A.

TITLE OF INVENTION: METHOD FOR GENERATING ACTIVATED T-CELLS
AND ANTIGEN-PULSED ANTIGEN-PRESENTING CELLS

NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESSES:
ADDRESSER: FOLEY & LARDNER
STREET: 3000 K Street, N.W.
CITY: Washington

STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/09/497,232
APPLICATION NUMBER: US/09/497,232
FILING DATE: 03-Feb-2000
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/09/074,410
FILING DATE: 08-MAY-1998
APPLICATION NUMBER: US 60/045,949
FILING DATE: 08-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Saxe, Bernhard D.
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 042881/0114
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5399
TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-497-232-10

Query Match 100.0%; Score 109; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVTSAADTRPAPGSTAPPAAH 20
DB 1 GVTSAADTRPAPGSTAPPAAH 20

RESULT 3
US-09-651-265-6
Sequence 6, Application US/09651265
Patent No. 6602660
GENERAL INFORMATION:
APPLICANT: AGRAMBL, BABITA
APPLICANT: LONGENECKER, B. MICHAEL
TITLE OF INVENTION: METHODS OF DETECTING T-CELL ACTIVATION
FILE REFERENCE: 042881/0151
CURRENT APPLICATION NUMBER: US/09/651,265
CURRENT FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 09/339,344
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: 60/090,916
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6
LENGTH: 20
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: peptide
US-09-651-265-6

Query Match 100.0%; Score 109; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVTSAADTRPAPGSTAPPAAH 20
DB 1 GVTSAADTRPAPGSTAPPAAH 20

RESULT 4

US-08-288-059-28
Sequence 28, Application US/08288059
Patent No. 5827666
GENERAL INFORMATION:
APPLICANT: FINN, OLIVERA J.
APPLICANT: FONTENOT, J. D.
APPLICANT: MONTELLARO, RONALD C.
TITLE OF INVENTION: SYNTHETIC MULTIPLE TANDEM REPEAT MUCIN
AND MUCIN-LIKE PEPTIDES, AND USES THEREOF
TITLE OF INVENTION: AND MUCIN-LIKE PEPTIDES, AND USES THEREOF
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/288,059
FILING DATE: 08-AUG-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: CHAPIN, MARILYN K.
REGISTRATION NUMBER: 35,843
REFERENCE/DOCKET NUMBER: 61137/205204
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3711
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-288-059-28

Query Match 100.0%; Score 109; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVTSAADTRPAPGSTAPPAAH 20
DB 1 GVTSAADTRPAPGSTAPPAAH 20

RESULT 5
US-08-134-198E-13
Sequence 13, Application US/08134198E
Patent No. 6150885
GENERAL INFORMATION:
APPLICANT: CANCER RESEARCH FUND
APPLICANT: OF CONTRA COSTA
APPLICANT: PETERSON, JERRY A.
APPLICANT: LARROCCA, DAVID J.
TITLE OF INVENTION: FUSION PROTEIN CONTAINING HMWG
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder & Poplowski
STREET: 444 South Flower Street, Suite 1900
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071

```

Query Match          100.0% ; Score 109 ; DB 9 ; Length 20 ;
Best Local Similarity 100.0% ; Pred. No. 1.5e-06 ;
Matches      20 ; Conservative      0 ; Mismatches      0 ; Indels      0 ; Gaps      0

CY      1 GGTSAADTRPAPGSTAPPAAH 20
        |||||
Db      1 GVTSAADTRPAPGSTAPPAAH 20

RESULT 2
US-09-984-333-6
; Sequence 6, Application US/09984333
; Patent No. US2002159569A1
; GENERAL INFORMATION:

```

APPLICANT: AGRAL, Babita
APPLICANT: LONGENECKER, Bryan Michael
APPLICANT: REDDISH, Mark Austin
TITLE OF INVENTION: SMALL PEPTIDE-BASED THERAPEUTICS FOR REVERSING
FILE REFERENCE: 042881/0116
CURRENT APPLICATION NUMBER: US/09/984,333
CURRENT FILING DATE: 2001-10-29
PRIOR APPLICATION NUMBER: US 09/182,887
PRIOR FILING DATE: 1998-10-30
PRIOR APPLICATION NUMBER: 60/064,146
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/065,209
PRIOR FILING DATE: 1997-11-12
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 20
TYPE: PRT
ORGANISM: Homo sapiens
US-09-984-333-6

Query Match 100.0%; Score 109; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVSADPTRPAPGSTAPPAAH 20
Db 1 GVSADPTRPAPGSTAPPAAH 20

RESULT 3
US-10-054-488-1
Sequence 1, Application US/10054488
Publication No. US20020142047A1
GENERAL INFORMATION:
APPLICANT: Mark E. Johnson
APPLICANT: Tricia Cecil
APPLICANT: Oliveira J. Finn
TITLE OF INVENTION: MICROSPHERE DELIVERY OF NOCIN PEPTIDES
FILE REFERENCE: 126.04USU1
CURRENT APPLICATION NUMBER: US/10/054,488
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: 60/262,699
PRIOR FILING DATE: 2001-01-19
NUMBER OF SEQ ID NOS: 1
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 20
TYPE: PRT
ORGANISM: Homo sapiens
US-10-054-488-1

Query Match 100.0%; Score 109; DB 13; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVSADPTRPAPGSTAPPAAH 20
Db 1 GVSADPTRPAPGSTAPPAAH 20

RESULT 4
US-09-996-069-9
Sequence 9, Application US/09996069
Publication No. US20030036199A1
GENERAL INFORMATION:
APPLICANT: Bamdad, Cynthia
APPLICANT: Bamdad, R. Shoshana
TITLE OF INVENTION: DIAGNOSTIC TUMOR MARKERS, DRUG SCREENING FOR TUMORIGENESIS INHIB
FILE REFERENCE: M01015/70071
CURRENT APPLICATION NUMBER: US/09/996,069

CURRENT FILING DATE: 2001-11-27
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.1
SEQ ID NO 9
LENGTH: 40
TYPE: PRT
ORGANISM: Homo sapiens
US-09-996-069-9

Query Match 100.0%; Score 109; DB 10; Length 40;
Best Local Similarity 100.0%; Pred. No. 3e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVSADPTRPAPGSTAPPAAH 20
Db 16 GVSADPTRPAPGSTAPPAAH 35

RESULT 5
US-10-106-876-14
Sequence 14, Application US/10106876
Publication No. US20030157160A1
GENERAL INFORMATION:
APPLICANT: BODZINSKI, WLADYSLAW A.
APPLICANT: KOGANTY, R. RAO
APPLICANT: KRANTZ, MARK J.
APPLICANT: LONGENECKER, B. MICHAEL
TITLE OF INVENTION: VACCINE FOR MODULATING BETWEEN T1 AND T2 IMMUNE
FILE REFERENCE: 042881-0176
CURRENT APPLICATION NUMBER: US/10/106,876
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: 60/278,698
PRIOR FILING DATE: 2001-03-27
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 14
LENGTH: 43
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-106-876-14

Query Match 100.0%; Score 109; DB 14; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.2e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVSADPTRPAPGSTAPPAAH 20
Db 19 GVSADPTRPAPGSTAPPAAH 38

RESULT 6
US-09-996-069-5
Sequence 5, Application US/09996069
Publication No. US20030036199A1
GENERAL INFORMATION:
APPLICANT: Bamdad, Cynthia
APPLICANT: Bamdad, R. Shoshana
TITLE OF INVENTION: DIAGNOSTIC TUMOR MARKERS, DRUG SCREENING FOR TUMORIGENESIS INHIB
FILE REFERENCE: M01015/70071
CURRENT APPLICATION NUMBER: US/09/996,069
CURRENT FILING DATE: 2001-11-27
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 46
TYPE: PRT
ORGANISM: Homo sapiens
US-09-996-069-5

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 04:52:45 ; Search time 21 Seconds

(without alignments)
91.611 Million cell updates/sec

Title: US-10-054-488-1

Perfect score: 109

Sequence: 1 GVTSAPDTPRAPGSTAPPAAH 20

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: PIR 78:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	109	100.0	347	2 S10571	mucin 1 precursor,
2	109	100.0	1344	1 A35175	mucin 1 precursor,
3	95	87.2	256	2 A60533	tumor-associated a
4	62	56.9	1334	2 T50568	probable multi-dom
5	57	52.3	428	2 T24769	hypothetical prote
6	56	51.4	267	2 F87665	hypothetical prote
7	56	51.4	569	2 A11347	hypothetical prote
8	55	51.4	2706	2 T28155	variant-specific s
9	55	50.5	539	2 G70520	probable cap prote
10	54	49.5	544	2 A71260	hypothetical prote
11	53	48.6	373	2 A70856	probable lppz prot
12	53	48.6	708	2 UC4364	gelatinase B (EC 3
13	53	48.6	708	2 S62907	gelatinase B (EC 3
14	52	47.7	115	2 F72570	hypothetical prote
15	52	47.7	182	2 S73046	pentacillin binding
16	52	47.7	196	2 T51838	blue copper binding
17	52	47.7	411	2 D86995	probable D-alanyl-
18	51.5	47.2	635	2 P75477	hypothetical prote
19	51	46.8	217	2 AC1887	hypothetical prote
20	51	46.8	382	2 T14336	RBD23 protein, iso
21	51	46.8	1201	2 A63007	hypothetical prote
22	50	45.9	402	2 S73773	dihydroliipoamide a
23	50	45.9	571	2 C75530	conserved hypocher
24	50	45.9	729	2 E70803	hypothetical prote
25	50	45.9	730	2 JC1456	gelatinase B (EC 3
26	50	45.9	1004	2 UC2221	major surface glyc
27	50	45.9	1173	1 A53430	1-phosphatidylinos
28	50	45.9	1199	2 JC4816	major surface glyc
29	50	45.9	1216	2 A28821	1-phosphatidylinos

30	50	45.9	3020	2 A43932	mucin 2 precursor,
31	49	45.0	78	2 T27876	hypothetical prote
32	49	45.0	175	2 T27875	hypothetical prote
33	49	45.0	430	2 JC2379	levanucrase (EC 2
34	49	45.0	447	2 T18264	cellulosome anchor
35	49	45.0	630	2 A39344	tumor-associated m
36	49	45.0	631	2 T22257	episiallin - mouse
37	49	45.0	1076	2 T30842	serine-repeat anti
38	49	45.0	1101	2 G70951	probable ATP-depen
39	49	45.0	1255	2 T31065	diaphanous protein
40	49	45.0	1315	2 A56101	collagen alpha 1(X
41	49	45.0	1774	2 B56101	collagen alpha 1(X
42	48.5	44.5	205	2 T34724	probable membrane
43	48.5	44.5	464	2 T36256	probable membrane
44	48.5	44.5	482	2 T36045	hypothetical prote
45	48	44.0	313	2 T33010	hypothetical prote

ALIGNMENTS

RESULT 1

S10571

mucin 1 precursor, secreted epithelial tumor antigen splice form - human

N/Contains: mucin 1 secreted breast-cancer-associated splice form

C/Species: Homo sapiens (man)

C/Date: 07-Apr-1994 #sequence revision 07-Apr-1994 #text_change 01-Dec-2000

C/Accession: S10571, UN0100, I56024, S09706, S10217

R/Wieschner, D.H.; Hareven, M.; Tsafaty, I.; Smorodinsky, N.; Horev, J.; Zaretsky, J.

Eur. J. Biochem. 189, 463-473, 1990

A/Title: Human epithelial tumor antigen CDNA sequences. Differential splicing may genera

A/Reference number: S10571, MUID:90276413, PMID:2351132

A/Accession: S10571

A/Molecule type: mRNA

A/Residues: 1-347 <MR>

A/Cross-references: EMBL:X52229, NID:936434, PIDN:CA36477.1, PID:936435

R/Tsafaty, I.; Hareven, M.; Horev, J.; Zaretsky, J.; Weiss, M.; Jeltsch, J.M.; Garnie

Gene 93, 313-318, 1990

A/Title: Isolation and characterization of an expressed hypervariable gene coding for a

A/Reference number: UN0100, MUID:91033045, PMID:1688329

A/Accession: UN0100

A/Status: Preliminary

A/Molecule type: DNA

A/Residues: 1-135, 'Q', 137-142, 'E', 144-163, 204-208, 'A', 210-347 <TSA>

A/Cross-references: GB:M5093, NID:9182252, PIDN:AA59612.1, PID:9182253

R/Xing, P.X.; Tjandra, J.J.; Reynolds, K.; McLaughlin, P.C.; Purcell, D.F.J.; McKenzie,

J. Immunol. 147, 3503-3509, 1990

A/Title: Reactivity of anti-human milk fat globule antibodies with synthetic peptides.

A/Reference number: I56024, MUID:89235154, PMID:2715633

A/Accession: I56024

A/Status: Preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 182-201 <RES>

A/Cross-references: GB:M26336, NID:9516622, PIDN:AA36336.1, PID:9516623

R/Tandler, S.J.B

Biochem. J. 267, 733-737, 1990

A/Title: Elements of secondary structure in a human epithelial mucin core peptide fragme

A/Reference number: S09706, MUID:90253387, PMID:2339983

A/Accession: S09706

A/Molecule type: protein

A/Residues: 182-201 <TEN>

C/Genetics: GDB:MUC1, PUM

A/Gene: GDB:MUC1, PUM

A/Cross-references: GDB:120705, OMIM:158340

A/Map position: 1q21-1q23

C/Keywords: alternative splicing; tandem repeat

P.1-23/Domain: signal sequence #status predicted <SIG>

P.24-347/Product: mucin 1, secreted epithelial tumor antigen splice form #status predict

P.24-163,204-347/Product: mucin 1, secreted breast-cancer-associated splice form #status

Query Match

Best Local Similarity 100.0%; Score 109; DB 2; Length 347;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 04:51:15 ; Search time 12 Seconds

(without alignments)
86.784 Million cell updates/sec

Title: US-10-054-488-1

Sequence: 1 GVTSAPDTRPAPGSTAPPAN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	109	100.0	1255	1 MUC1_HUMAN	P15941 h mucin 1 p
2	91	83.5	475	1 MUC1_HYDRA	Q23433 hylobates p
3	54	49.5	544	1 Y366_TREEA	O81933 treponema p
4	54	49.5	564	1 ZYX_MOUSE	O62523 mus musculu
5	53	48.6	708	1 MM09_RAT	P50282 rattus norv
6	52	47.7	196	1 BCPI_ARATH	O07488 arabidopsis
7	52	47.7	333	1 NK32_HUMAN	P78367 homo sapien
8	51	46.8	303	1 PCPR_SPHOR	P56678 sphingobium
9	50	45.9	402	1 ODP2_MYCEN	P75392 mycoplasma
10	50	45.9	704	1 MM09_CANFA	O18733 canis fami
11	50	45.9	1125	1 IF2_PROMM	Q745M4 prochloroco
12	50	45.9	1216	1 PIR1_MOUSE	O941b3 mus musculu
13	50	45.9	1216	1 PIR1_MOUSE	O43711 homo sapien
14	49	45.0	291	1 TLX3_HUMAN	O55144 mus musculu
15	49	45.0	291	1 TLX3_HUMAN	O93367 gallus galli
16	49	45.0	297	1 TLX3_MOUSE	O06846 clostridium
17	49	45.0	447	1 ANCA_CLOTH	Q04296 mus musculu
18	49	45.0	630	1 MUC1_MOUSE	O08808 mus musculu
19	49	45.0	1255	1 DIAL_MOUSE	P39061 mus musculu
20	49	45.0	1774	1 CALH_MOUSE	O29V48 rattus norv
21	49	45.0	2167	1 SHK1_RAT	O19525 homo sapien
22	48	44.0	351	1 FX11_HUMAN	O88506 rattus norv
23	48	44.0	553	1 SPK1_RAT	O88506 rattus norv
24	48	44.0	642	1 M95_MOUSE	O98011 mus musculu
25	48	44.0	670	1 SR72_CANFA	P33731 canis fami
26	48	44.0	670	1 SR72_CANFA	O76094 homo sapien
27	48	44.0	1229	1 P121_HUMAN	O94291 homo sapien
28	48	44.0	1496	1 CA25_HUMAN	O05927 homo sapien
29	48	44.0	1575	1 SYT1_HUMAN	O44426 homo sapien
30	47	43.1	398	1 ODO2_AZOVI	P20708 azotobacter
31	47	43.1	399	1 ALX4_MOUSE	O31213 mus musculu
32	47	43.1	473	1 RGRP_HUMAN	O81234 homo sapien
33	47	43.1	522	1 KAPR_USTWA	P49605 ustiliago ma

34	47	43.1	619	1 LCPI_MOUSE	O8b11 mus musculu
35	47	43.1	619	1 LCPI_RAT	O99pm1 rattus norv
36	47	43.1	788	1 TRS1_HCVMA	P09698 human cytom
37	47	43.1	846	1 TRS1_HCVMA	P09715 human cytom
38	47	43.1	865	1 CPN_DROME	Q02910 drosophila
39	47	43.1	901	1 A180_MOUSE	O61548 mus musculu
40	47	43.1	915	1 A180_RAT	O05140 rattus norv
41	47	43.1	4391	1 PGBM_HUMAN	P98160 homo sapien
42	46.5	42.7	553	1 FXC1_HUMAN	O12948 homo sapien
43	46.5	42.7	1827	1 MAP2_HUMAN	P11137 homo sapien
44	46.5	42.7	5038	1 PCLO_MOUSE	O8qy77 mus musculu
45	46	42.2	145	1 COB8_HUMAN	P59052 homo sapien

ALIGNMENTS

RESULT 1			
ID	MUC1_HUMAN	STANDARD:	PRT: 1255 AA.
AC	P15941; P15931; P15942; P17626; Q14128; Q14876; Q16437; Q16442; Q16615; Q9BX44; Q9UE75; Q9UE76; Q9UOL1; Q9Y4J2;		
DT	01-JAN-1990 (Rel. 13, Created)		
DT	01-APR-1990 (Rel. 14, Last sequence update)		
DT	10-OCT-2003 (Rel. 42, Last annotation update)		
DE	Mucin 1 precursor (MUC-1) (Polymorphic epithelial mucin) (PEM) (PEMT)		
DE	(Epithelial) (Tumor-associated mucin) (Carcinoma-associated mucin)		
DE	(Tumor-associated epithelial membrane antigen) (EMA) (H23AG) (Peanut		
DE	reactive urinary mucin) (PUM) (Breast carcinoma-associated antigen		
DE	DF3) (CD227 antigen).		
GN	MUC1.		
OC	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	(1)		
RP	SEQUENCE FROM N.A. (ISOFORM 1).		
RC	TISSUE=pancreas;		
RX	MEDLINE=90368716; PubMed=2394722;		
RA	Ian M.S., Barra S.K., Qi W.-N., Metzgar R.S., Hollingsworth M.A.;		
RT	"Cloning and sequencing of a human pancreatic tumor mucin cDNA.";		
RL	J. Biol. Chem. 265:15294-15299 (1990).		
RN	(2)		
RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).		
RX	MEDLINE=90202794; PubMed=2318825;		
RA	Ligendberg M.J.L., Vos H.L., Gennissen A.M.C., Hilkens J.;		
RT	"Epithelial, a carcinoma-associated mucin, is generated by a		
RT	polymorphic gene encoding splice variants with alternative amino		
RT	terminal.";		
RL	J. Biol. Chem. 265:5573-5578 (1990).		
RN	(3)		
RP	SEQUENCE FROM N.A. (ISOFORM 1).		
RC	TISSUE=Breast carcinoma;		
RX	MEDLINE=90368715; PubMed=1697589;		
RA	Gendler S.U., Lancaster C.A., Taylor-Papadimitriou J., Duhig T.;		
RA	Peat N., Burchell J., Pemberton L., Lalani E.-N., Wilson D.;		
RT	"Molecular cloning and expression of human tumor-associated		
RT	polymorphic epithelial mucin.";		
RL	J. Biol. Chem. 265:15286-15293 (1990).		
RN	(4)		
RP	SEQUENCE FROM N.A. (ISOFORM 1).		
RX	MEDLINE=91097524; PubMed=2268309;		
RA	Lancaster C.A., Peat N., Duhig T., Wilson D.;		
RA	Taylor-Papadimitriou J., Gendler S.U.;		
RT	"Structure and expression of the human polymorphic epithelial mucin		
RT	gene: an expressed VNTR unit.";		
RL	Biochem. Biophys. Res. Commun. 173:1019-1029 (1990).		
RN	(5)		
RP	SEQUENCE FROM N.A. (ISOFORM 5).		
RC	TISSUE=Breast carcinoma;		
RX	MEDLINE=90276413; PubMed=2351132;		
RA	Wreschner D.H., Hareven M., Tsarfaty I., Smorodinsky N., Horev J.;		
RA	Zaretsky J., Kotkes P., Weiss M., Lathe R., Dion A., Keydar I.;		

RT "Human epithelial tumor antigen cDNA sequences. Differential splicing
RT may generate multiple protein forms.";
RL Eur. J. Biochem. 189:463-473(1990).
RN [61]
RP SEQUENCE FROM N.A.
RC TISSUE-Breast carcinoma;
RX MEDLINE=90276414; PubMed=2112460;
RA Hareven M., Tsarfaty I., Zaretsky J., Kotkes P., Horev J.,
RA Zrihan S., Weiss M., Green S., Lathe R., Keydar I., Wreschner D.H.;
RT "A transcribed gene, containing a variable number of tandem repeats,
RT codes for a human epithelial tumor antigen. cDNA cloning, expression
RT of the transcribed gene and over-expression in breast cancer
RT tissue.";
RL Eur. J. Biochem. 189:475-486(1990).
RN [7]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=91033045; PubMed=1688329;
RA Tsarfaty I., Hareven M., Horev J., Zaretsky J., Weiss M.,
RA Jeltsch J.M., Garnier J.M., Lathe R., Keydar I., Wreschner D.H.;
RT "Isolation and characterization of an expressed hypervariable gene
RT coding for a breast-cancer-associated antigen.";
RL Gene 93:313-318(1990).
RN [8]
RP SEQUENCE FROM N.A. (ISOFORM 7).
RX MEDLINE=95010060; PubMed=7925397;
RA Zrihan-Licht S., Vos H.L., Baruch A., Elroy-Stein O., Sagiv D.,
RA Keydar I., Halkens J., Wreschner D.H.;
RT "Characterization and molecular cloning of a novel MUC1 protein,
RT devoid of tandem repeats, expressed in human breast cancer tissue.";
RL Eur. J. Biochem. 224:787-795(1994).
RN [9]
RP SEQUENCE FROM N.A. (ISOFORMS 6; 7 AND 8).
RX MEDLINE=97355747; PubMed=9212228;
RA Oosterkamp H.M., Scheiner L., Stefanova M.C., Lloyd K.O.,
RA Finsied C.L.;
RT "Comparison of MUC-1 mucin expression in epithelial and non-epithelial
RT cancer cell lines and demonstration of a new short variant form
RT (MUC-1/2).";
RL Int. J. Cancer 72:87-94(1997).
RN [10]
RP SEQUENCE FROM N.A. (ISOFORM 7).
RX Zhang L.X., Li C.H.;
RT "Molecular cloning of an isoform of MUC1, MUC1/Y.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [11]
RP SEQUENCE FROM N.A. (ISOFORM 9).
RC TISSUE-Epithelial cancer;
RX Zhang L.X., Li C.H., Sun L.Y., Yue W.;
RT "Cloning of a new potential secreted short variant form of MUC1 mucin
RT in epithelial cancer cell line.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [12]
RP PARTIAL SEQUENCE FROM N.A.
RX MEDLINE=88330762; PubMed=3417635;
RA Gendler S.J., Taylor-Papadimitriou J., Duhig T., Reichard J.,
RA Burchell J.;
RT "A highly immunogenic region of a human polymorphic epithelial mucin
RT expressed by carcinomas is made up of tandem repeats.";
RL J. Biol. Chem. 263:12820-12823(1988).
RN [13]
RP SEQUENCE OF 1-160 FROM N.A. (ISOFORM 2).
RX MEDLINE=90088473; PubMed=2597151;
RA Abe M., Siddiqui J., Kufe D.;
RT "Sequence analysis of the 5' region of the human DP3 breast
RT carcinoma-associated antigen gene.";
RL Biochem. Biophys. Res. Commun. 165:644-649(1989).
RN [14]
RP SEQUENCE OF 1-109 FROM N.A. (ISOFORM 2).
RC TISSUE-Thyroid;
RX MEDLINE=96183746; PubMed=8608966;
RA Weiss M., Baruch A., Keydar I., Wreschner D.H.;
RT "Preoperative diagnosis of thyroid papillary carcinoma by reverse
RT transcriptase polymerase chain reaction of the MUC1 gene.";

RL Int. J. Cancer 66:55-59(1996).
RN [15]
RP SEQUENCE OF 1-89 FROM N.A.
RC TISSUE-Lung;
RX MEDLINE=96181716; PubMed=8604237;
RA Yu C.J., Yang P.C., Shew P.Y., Hong T.M., Yang S.C., Lee Y.C.,
RA Lee L.N., Luh K.T., Wu C.W.;
RT "Mucin mRNA expression in lung adenocarcinoma cell lines and
RT tissues.";
RL Oncology 53:118-126(1996).
RN [16]
RP SEQUENCE OF 1-46 FROM N.A. (ISOFORMS 3 AND 4).
RC TISSUE-Breast carcinoma;
RA Butwela L., Liu Q., Lugmani Y.A., Gomm J.J., Coombes R.C.;
RL Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.
RN [17]
RP CARBOHYDRATE-LINKAGE SITES WITHIN THE REPEAT.
RX MEDLINE=97460054; PubMed=9312074;
RA Mueller S., Goletz S., Packer N., Goolley A.A., Lawson A.M.,
RA Hanisch F.-G.;
RT "Localization of O-glycosylation sites on glycopeptide fragments from
RT lactation-associated MUC1. All putative sites within the tandem
RT repeat are glycosylation targets in vivo.";
RL J. Biol. Chem. 272:24780-24793(1997).
RN [18]
RP CARBOHYDRATE-LINKAGE SITES WITHIN THE REPEAT.
RX MEDLINE=99303572; PubMed=10373415;
RA Mueller S., Alving K., Peter-Katalinic J., Zachara N., Goolley A.A.,
RA Hanisch F.-G.;
RT "High density O-glycosylation on tandem repeat peptide from secretory
RT MUC1 of T47D breast cancer cells.";
RL J. Biol. Chem. 274:18165-18172(1999).
RN [19]
RP POLYMORPHISM WITHIN THE REPEAT.
RX MEDLINE=21359366; PubMed=1150974;
RA Engelmann K., Baldus S.B., Hanisch F.-G.;
RT "Identification and topology of variant sequences within individual
RT repeat domains of the human epithelial tumor mucin MUC1.";
RL J. Biol. Chem. 276:27764-27769(2001).
RN [20]
RP CHARACTERIZATION OF ISOFORM Y, AND MUTAGENESIS OF ASP-1116.
RX MEDLINE=99211485; PubMed=10197628;
RA Baruch A., Hartmann M.-L., Yoeli M., Adereth Y., Greenstein S.,
RA Stadler Y., Skornik Y., Zaretsky J., Smorodinsky N.I., Keydar I.,
RA Wreschner D.H.;
RT "The breast cancer-associated MUC1 gene generates both a receptor and
RT its cognate binding protein.";
RL Cancer Res. 59:1552-1561(1999).
RN [21]
RP PARTIAL SEQUENCE, AND CHARACTERIZATION OF CLEAVAGE SITE.
RX MEDLINE=21240104; PubMed=11341784;
RA Parry S., Silverman H.S., McDermott K., Willis A., Hollingsworth M.A.,
RA Harris A.;
RT "Identification of MUC1 proteolytic cleavage sites in vivo.";
RL Biochem. Biophys. Res. Commun. 283:715-720(2001).
RN [22]
RP CHARACTERIZATION.
RX MEDLINE=21836452; PubMed=11847293;
RA Wreschner D.H., McGuckin M.A., Williams S.J., Baruch A., Yoeli M.,
RA Ziv R., Okun L., Zaretsky J., Smorodinsky N., Keydar I., Neophytou P.,
RA Stacey M., Lin H.-H., Gordon S.;
RT "Generation of ligand-receptor alliances by 'SEA' module-mediated
RT cleavage of membrane-associated mucin proteins.";
RL Protein Sci. 11:698-706(2002).
RN [23]
RP PHOSPHORYLATION.
RX MEDLINE=95080414; PubMed=7988707;
RA Zrihan-Licht S., Baruch A., Elroy-Stein O., Keydar I., Wreschner D.H.;
RT "Tyrosine phosphorylation of the MUC1 breast cancer membrane proteins.
RT Cytokine receptor-like molecules.";
RL FEBS Lett. 356:130-136(1994).
CC -!- FUNCTION: May play a role in adhesive functions and in cell-cell
interactions, metastasis and signaling. May provide a protective

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 04:53:15 ; Search time 40 Seconds

(without alignments)
157.759 Million cell updates/sec

Title: US-10-054-488-1

Sequence: 1 GVTSPAPDPAPGAPSTAPPAN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_25: *
1: sp.archaea: *
2: sp.bacteria: *
3: sp.fungi: *
4: sp.human: *
5: sp.invertebrate: *
6: sp.mammal: *
7: sp.mhc: *
8: sp.organelle: *
9: sp.phage: *
10: sp.plant: *
11: sp.rodent: *
12: sp.virus: *
13: sp.vertebrate: *
14: sp.unclassified: *
15: sp.virus: *
16: sp.bacteriophage: *
17: sp.archae: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86	78.9	20	4 Q9UM18	Q9um18 homo sapien
2	77	70.6	553	6 Q9MZL1	Q9mz11 macaca mula
3	63	57.8	745	16 Q89X06	Q89x06 bradyrhizob
4	62	56.9	1334	10 Q8LIP8	Q8lip8 oryza sativ
5	62	56.9	1334	16 Q9RKR9	Q9rkr9 streptomyce
6	60	55.0	162	6 Q8MUM4	Q8mum4 equus hemio
7	60	55.0	162	6 Q8MUM4	Q8mum4 equus asinu
8	60	55.0	168	6 Q8MUM4	Q8mum4 equus grevy
9	60	55.0	168	6 Q8MUM4	Q8mum4 equus zebra
10	60	55.0	705	2 Q848C3	Q848c3 streptomyce
11	59	54.1	360	16 Q8P7V2	Q8p7v2 xanthomonas
12	59	54.1	889	16 Q9F2N5	Q9f2n5 streptomyce
13	57	52.3	428	5 Q22369	Q22369 caenorhabdi
14	56	51.4	267	16 Q9A346	Q9a346 caulobacter
15	56	51.4	569	16 Q9KGV9	Q9kgv9 listeria mo
16	56	51.4	2706	5 Q15870	Q15870 plasmodium

17	55	50.5	329	10 Q9GM15	Q9gm15 zea mays (m
18	55	50.5	539	16 Q86334	Q86334 mycobacteri
19	55	50.5	539	16 Q7VWM2	Q7vwm2 mycobacteri
20	54	49.5	433	4 Q9UMH8	Q9umh8 homo sapien
21	54	49.5	481	4 Q9NZ07	Q9nzk2 homo sapien
22	54	49.5	484	4 Q9UMX1	Q9umx1 homo sapien
23	54	49.5	484	4 Q9UMH2	Q9umh2 homo sapien
24	54	49.5	508	11 Q8CD55	Q8cd55 mus muscu
25	54	49.5	533	11 Q7Q0E2	Q7q0e2 mus muscu
26	53	48.6	168	6 Q8MUM7	Q8mum7 equus cabal
27	53	48.6	373	16 Q53253	Q53253 mycobacteri
28	53	48.6	373	16 Q7XG5	Q7xg5 mycobacteri
29	53	48.6	463	10 Q948F9	Q948f9 oryza sativ
30	53	48.6	505	15 Q8JAX4	Q8jax4 human immun
31	53	48.6	506	15 Q8JAX3	Q8jax3 human immun
32	53	48.6	516	5 Q9XYH5	Q9xyh5 strongyloce
33	53	48.6	554	6 Q95L89	Q95l89 bos taurus
34	53	48.6	580	6 Q8MUM4	Q8mum4 bos taurus
35	53	48.6	725	4 Q8NDJ6	Q8ndj6 homo sapien
36	52	47.7	115	17 Q9YAU7	Q9yau7 aeropyrum p
37	52	47.7	162	6 Q8MUM6	Q8mum6 equus cabal
38	52	47.7	168	6 Q8MUM5	Q8mum5 equus cabal
39	52	47.7	182	2 Q49921	Q49921 mycobacteri
40	52	47.7	186	11 Q8C4T0	Q8c4t0 mus muscu
41	52	47.7	252	10 Q7XIF6	Q7xif6 oryza sativ
42	52	47.7	333	10 Q8SAY7	Q8say7 oryza sativ
43	52	47.7	333	10 Q850X2	Q850x2 oryza sativ
44	52	47.7	411	16 Q9CCM2	Q9ccm2 mycobacteri
45	52	47.7	440	16 Q9EX20	Q9ex20 streptomyce

ALIGNMENTS

RESULT 1

Q9UM18 PRELIMINARY; PRT; 20 AA.

AC Q9UM18; 01-MAY-2000 (TREMELREL. 13, Created)

DT 01-MAY-2000 (TREMELREL. 13, Last sequence update)

DT 01-JUN-2003 (TREMELREL. 24, Last annotation update)

DE Mucin (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_Taxid=9606;

RN [1]

RP SEQUENCE FROM N.A.

RY MEDLINE=89235154; PubMed=2715633;

RA Xing P.X., Tjandra U.J., Reynolds K., McLaughlin P.J., Purcell D.F.J.,

RT "Reactivity of anti-human milk fat globule antibodies with synthetic

peptides."

RL EMBL; M26316; AAA6363.1; -

DR PIR; S10571; S10571.

FT NON_TER

FT NON_TER

SQ SEQUENCE

20 AA; 1887 MW; 5B3473EABEAPAD87 CRC64;

Query Match 78.9%; Score 86; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 PDTRPAPGAPSTAPPAN 20

DB 1 PDTRPAPGAPSTAPPAN 15

RESULT 2

Q9MZL1 PRELIMINARY; PRT; 553 AA.

AC Q9MZL1;

```

DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Mucin 1 (Fragment).
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20330533; PubMed=10869775;
RA Vaughan H.A.; Ho D.W.M.; Karanikas V.; Sandrin M.S.; McKenzie I.F.C.;
RA Pleteraz G.A.;
RT "The Immune Response of Mice and Cynomolgus Monkeys to Macaque Mucin-1
  Mannan."
RL Vaccine 18:3297-3309(2000).
DR EMBL: AF176847; AAP82403.1;
DR InterPro: IPR001064; Crystalin.
DR InterPro: IPR000082; SEA_domain.
DR Pfam: PF01350; SEA; 1.
DR SMART: SM00200; SEA; 1.
DR PROSITE: PS00225; CRYSTALLIN_BETAGAMMA; 1.
DR PROSITE: PS50024; SEA; 1.
FT NON_TER 1
FT NON_TER 553
SQ SEQUENCE 553 AA; 55778 MW; 6D7E6DD2E929318 CRC64;

Query Match
Best Local Similarity 83.3%; Score 77; DB 6; Length 553;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 VTSPDTRPAPGSTAPPA 19
DB 76 VTSPDTRPAPGSTAPPA 93

RESULT 3
Q89X06 PRELIMINARY; PRT; 745 AA.
ID Q89X06;
AC Q89X06;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE BIR0521 protein.
GN BIR0521.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneke T.; Nakamura Y.; Sato S.; Minamisawa K.; Uchiumi T.;
RA Sasamoto S.; Matsumoto A.; Idesawa K.; Iriyuchi M.; Kawashima K.;
RA Kohara M.; Matsumoto M.; Shingo S.; Tsuruoka H.; Wada T.; Yamada M.;
RA Tabata S.; genomic sequence of nitrogen-fixing symbiotic bacterium
RT "Bradyrhizobium japonicum USDA110."
RL DNA Res. 9:189-197(2002).
DR EMBL: AP005936; BAC45786.1; -.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0004672; F:protein kinase activity; IEA.
DR GO: GO:0005199; F:structural constituent of cell wall; IEA.
DR GO: GO:0006468; F:protein amino acid phosphorylation; IEA.
DR InterPro: IPR006665; OmpA/MotB.
DR InterPro: IPR003882; PstII_extensin.
DR InterPro: IPR007119; Prot_kinase.
DR Pfam: PF00691; OmpA; 1.
DR PRINTS: PRO1218; PSTLEXTENSIN.
DR ProDom: PD000001; Prot_kinase; 1.
KW Complete proteome.

```

```

SQ SEQUENCE 745 AA; 74545 MW; 155EDFC74DBCF6D CRC64;

Query Match
Best Local Similarity 66.7%; Score 63; DB 16; Length 745;
Matches 12; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GVTSPDTRPAPGSTAPP 18
DB 274 GATPAPTTTPAGGTATP 291

RESULT 4
Q8LIP8 PRELIMINARY; PRT; 564 AA.
ID Q8LIP8;
AC Q8LIP8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative PR1-interacting factor G.
GN OJ1370_E02.13 OR OJ1354_H07.23.
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Nipponbare;
RA Sasaki T.; Matsumoto T.; Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
  clone:OJ1370_E02."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP003755; BAC21341.1; -.
DR EMBL: AP003756; BAC10056.1; -.
DR Gramene: O8LIP8; -.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003677; F:DNA binding; IEA.
DR InterPro: IPR01005; Myo_DNA_binding.
DR PROSITE: PS00037; MYB_1; 1.
SQ SEQUENCE 564 AA; 61268 MW; F6D89CF602B5ADF0 CRC64;

Query Match
Best Local Similarity 63.2%; Score 62; DB 10; Length 564;
Matches 12; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 GVTSPDTRPAPGSTAPPA 19
DB 34 GVTSPDTRPAPGSTAPPA 52

RESULT 5
Q8RKR9 PRELIMINARY; PRT; 1334 AA.
ID Q8RKR9;
AC Q8RKR9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative multi-domain regulatory protein.
GN SC02259 OR SCC75A.05C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=A3 (2);

```